Journal of Educational Psychology
2010, Vol. 102, No. 3, 687–700

0022-0663/10/$12.00 DOI: 10.1037/a0019452

Reading Comprehension of Scientific Text: A Domain-Specific Test of the Direct and Inferential Mediation Model of Reading Comprehension

Jennifer G. Cromley and Lindsey E. Snyder-Hogan
Temple University

Ulana A. Luciw-Dubas
National Board of Medical Examiners, Philadelphia, Pennsylvania

Reading comprehension is strongly associated with academic achievement, including science achievement. A better understanding of reading comprehension processes in science text might hold promise for improving science achievement in the long run. We tested the fit of the direct and inferential mediation (DIME) model of reading comprehension (Cromley & Azevedo, 2007) with 737 students in an introductory biology course required for majors. Participants completed multiple choice measures of biology-specific prior-topic knowledge, inference, reading strategy use, reading vocabulary, word reading fluency, and reading comprehension in small groups in our laboratory. Using structural equation modeling to test the fit of the DIME model to the data, we found excellent fit indices for all models. However, the original DIME model fit significantly worse than the measurement model, and a modified model that included a path from reading vocabulary to reading strategy use fit significantly better. Results from the modified model suggest that comprehension interventions for undergraduate students with biology majors might use preteaching to build topic knowledge. We discuss the need for future experimental studies to confirm the vocabulary-reading strategies link.

Keywords: comprehension, knowledge, strategies, inference

Reading comprehension is a critical enabling skill for school success, and comprehension test scores show medium-to-high correlations with academic achievement (correlations from .42 to .66; Law, Chan, & Sachs, 2008; Savolainen, Ahonen, Aro, Tolvanen, & Holopainen, 2008). Comprehension is a complex phenomenon and is the result of the combined effects of several subcomponent variables. Although researchers are not in complete agreement on which skills and knowledge are required for comprehension, several variables are included in numerous models and theories of reading comprehension: inference, reading comprehension strategies, reading vocabulary, word reading, and prior topic knowledge (see Kintsch, 1998; Nation & Angell, 2006; O'Reilly & McNamara, 2007a, 2007b; Paris & Stahl, 2005; Perfetti, Yang, & Schmalhofer, 2008; van den Broek & Kendeou, 2008).

There is great concern about underperformance in science for students in the United States compared with other countries (National Research Council, 2007). Reading comprehension of scientific text may be critical for science achievement (Otero, León, & Graesser, 2002; Ozuru, Dempsey, & McNamara, 2009); a small body of research shows impressively large correlations between reading comprehension and science proficiency (Cromley, 2009; O'Reilly & McNamara, 2007a). A better understanding of reading comprehension in science text might hold promise for improving science achievement in the long run.

Research on various predictors of reading comprehension in the years beyond beginning reading acquisition is now a relatively large and mature field, although there are several controversies that have not been resolved. We recently synthesized this research to create a new model of reading comprehension—the direct and inferential mediation (DIME; Cromley & Azevedo, 2007) model—and we validated the model with samples of high school and undergraduate students (Cromley & Snyder, 2007). A major strength of the DIME model is that paths are only included in the model when they meet the strict criterion of being based on one or more experimental studies published in a peer-reviewed journal. The model, however, has several weaknesses: (a) It is domain-general, but as students progress through college, their reading tends to focus on a single domain—would the DIME model hold equally well with science-major students reading scientific text?; (b) relatively small samples were used to test the fit of the model; and (c) the model did not show an ideal fit to the data.

First, we are interested in comprehension of scientific text, which is characterized by explanatory text structure and heavy vocabulary demands (Otero et al., 2002). Does the DIME model hold equally well with this specific type of text? If so, this will be evidence that the model can be generalized beyond domain-general text to the single domain of biology. Second, we wish to validate this new theoretical model with a large sample. Prior findings could have been due to sampling error, whereas we can be more confident in results from a larger sample. Finding a good fit

Jennifer G. Cromley and Lindsey E. Snyder-Hogan, Department of Psychological Studies in Education, Temple University; Ulana A. Luciw-Dubas, National Board of Medical Examiners, Philadelphia, Pennsylvania.

Previous versions of this research were presented at the 2007 annual conference of the American Educational Research Association (Chicago, Illinois) and the 2008 annual conference of the Society for the Scientific Study of Reading (Asheville, North Carolina). This research was supported by funds from the 2006–2007 Temple University Return of Overhead Research Incentive Grant Program.

Correspondence concerning this article should be addressed to Jennifer G. Cromley, Department of Psychological Studies in Education, Temple University, Ritter Annex 201, 1301 Cecil B. Moore Avenue, Philadelphia, PA 19122-6091. E-mail: jcromley@temple.edu

of the model with a larger sample would provide further evidence of the generality of the model beyond the small sample of high school students reading domain-general text. Third, the fit of the model to the data in our original test of the model (Cromley & Azevedo, 2007) was marginal; could the fit be improved by further refining the model and by using more sophisticated analytical techniques?

In the current study, we test a domain-specific version of the DIME model with a larger sample, use structural equation modeling (SEM) to test the fit of the original model, and test several variations of the model in an attempt to improve the fit of the model. These results could provide further evidence for the generalizability of the DIME model beyond domain-general text to biology text and could lead to approaches for improving comprehension among science students.

Recent Studies Supporting the DIME Model

The DIME model (Cromley & Azevedo, 2007) builds on evidence from 98 published experimental studies conducted with students in fourth grade and above to support the various paths (see Figure 1). We refer interested readers to that in-depth literature review rather than duplicating it here. We augment that review with articles published since 2006 that meet the same criteria and show continued support for the hypothesized paths in the DIME model. Below, we briefly review recent experimental studies relevant to the model.

Inference

Authors do not make explicit the relationships among every proposition in text. A variety of inferences—at the inter- and intrasentence level—are therefore key for comprehension. Inferences typically found in reading include conclusions drawn between propositions in text, either within sentences or between adjacent sentences (bridging inferences), and conclusions drawn between prior topic knowledge and the current text (elaborative inferences). Training in drawing inferences in text increases comprehension (e.g., Best, Rowe, Ozuru, & McNamara, 2005; McKeown, Beck, & Blake, 2009; O'Reilly & McNamara, 2007b), as does making text more coherent by adding explicit referential connections to increase inferential processing (Ainsworth & Burcham, 2007; Gilabert, Martinez, & Vidal-Abarca, 2005) and by encouraging readers to self-explain (Ainsworth & Burcham, 2007; Best et al., 2005).

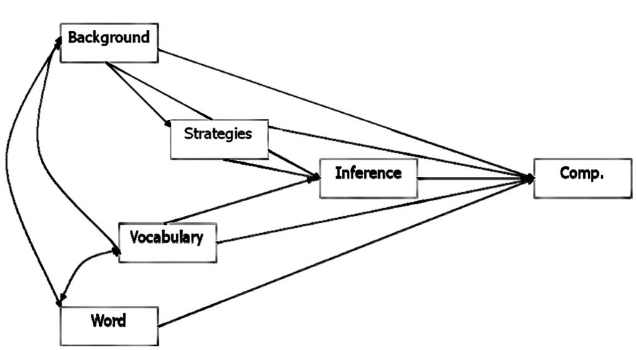

Figure 1. The original direct and inferential mediation model (Cromley & Azevedo, 2007). Comp. = comprehension.

Increasing the number and quality of inferences drawn while reading science text increases comprehension. For example, Self-Explanation Reading Training (SERT) is a program that teaches readers "to draw deeper inferences based on the active, constructive processing of connections" (Best et al., 2005, p. 77). Research with middle school through undergraduate students—using SERT and its computer-based successor called iSTART—in regards to reading scientific text suggests that by teaching readers to generate and elaborate on inferences that connect new information to previous sections of text, a deeper level of comprehension can be achieved. McKeown et al. (2009) described the results of their 2-year study in which students were taught content-related strategies while reading narrative and scientific texts. Results of this study suggest that elementary-age students who are taught to ask "open, meaning-based questions" (McKeown et al., 2009, p. 218) while reading show increased scores on recall measures that transfer to standardized reading comprehension measures.

A second line of evidence regarding the importance of inference concerns changing text to include more explicit connections between segments (i.e., to make the text more coherent). Text that is made more coherent by adding explicit referential connections increases readers' inferential processing; this in turn this leads to better reading comprehension. For example, O'Reilly and McNamara (2007b) provided low- or high-cohesion science texts to undergraduate students with low prior knowledge. Bridging and elaborative inference questions were used as a measure of reading comprehension. Results suggest that low-knowledge readers benefit from high-cohesion text because they are able to find cues that lead to increases in their comprehension. Likewise, Gilabert et al. (2005) presented low- and high-knowledge eighth-grade participants with one of three versions of a history passage: the original, a more cohesive version aimed at promoting reader's inferential activity, and a less cohesive version aimed at reducing inferences while reading. The more cohesive version had significant, positive effects on recall and inference measures. These results suggest that explicit connections made within text are associated with higher levels of inference.

A third approach to increasing inferential activity while reading is training readers to self-explain. For example, Ainsworth and Burcham (2007) trained some undergraduate participants to self-explain while reading. All participants were given either a minimally or maximally cohesive science passage as part of the study. Results suggest that the trained participants self-explained more while reading the text; these effects on comprehension were above and beyond those due to the high-cohesion text manipulation. In SERT, individuals are taught a set of self-explanation reading comprehension strategies and are shown how to practice using these strategies while reading science text (Best et al., 2005). The SERT and iSTART programs have been shown to increase the quality of self-explanation and comprehension of science texts (Best et al., 2005; McNamara, 2004).

Reading Comprehension Strategies

There is a large body of literature on teaching reading comprehension strategies, such as summarizing, making concept maps, questioning, and comprehension monitoring. These strategies have

typically been termed cognitive strategies, metacognitive strategies, or self-regulatory strategies in the reading comprehension literature. Recent research continues to suggest that instruction in the use of these strategies generally—but not universally—is associated with increases in reading comprehension in undergraduate (e.g., Callender & McDaniel, 2007; Gunn, 2008; McCrudden, Schraw, & Lehman, 2009), middle school (e.g., Lau & Chan, 2007; Proctor, Dalton, & Grisham, 2007; Souvignier & Mokhlesgerami, 2006; Sung, Chang, & Huang, 2008), and elementary school (e.g., Kim & White, 2008; Spörer, Brunstein, & Kieschke, 2009; Wigfield et al., 2008) students. Strategies taught across these 10 studies include generation, concept maps, reciprocal teaching, summarizing, monitoring, text structure, imagery, highlighting, re-reading, questioning, sentence combining, and concept-oriented reading instruction. Samples in these studies included regular readers, struggling readers, and English language learners; were conducted across several countries; and included students at different levels of various individual difference variables, such as prior knowledge and motivation. Although earlier research on strategy instruction was more pessimistic about achieving transfer of instructed strategies (Hattie, Biggs, & Purdie, 1996; Salomon & Perkins, 1989), several recent studies have shown evidence of transfer of instructed reading comprehension strategies for text structure (Broer, Aarnoutse, Kieviet, & van Leeuwe, 2002) and for multiple-strategy instruction programs (De Corte, Verschaffel, & Van De Ven, 2001; Souvignier & Mokhlesgerami, 2006).

Teaching specific reading comprehension strategies to students also increases correct inferences in narrative text (Gunn, 2008; McDaniel, Howard, & Einstein, 2009; Ozgungor & Guthrie, 2004; Spörer et al., 2009; Sung et al., 2008). Gunn (2008) provided brief training to a subset of 63 students with undergraduate psychology majors in generating questions about their reading, including memory-based questions and critical thinking questions. After reading a science passage, the students in this condition scored higher on a measure of elaborative and bridging inferences than students in unstructured questioning or control conditions. Similarly, Ozgungor and Guthrie (2004) taught a subset of 119 students with psychology and biology majors to ask themselves elaborative questions, which meld new and prior knowledge, as they read through an essay about a neuropsychological phenomenon. The students in the elaborative condition made more accurate inferences, particularly when they were more interested in the material. In another study, a subset of 72 students with undergraduate psychology majors were taught to read, recite from memory, and then review educational texts (McDaniel et al., 2009). These students performed significantly better on short-answer and multiple-choice tests that measured inferences compared with students in re-reading and note-taking conditions. These same results were found after a 1-week delay.

Two recent studies of German and Taiwanese upper elementary school students have additionally found positive effects of strategy instruction on inference (Spörer et al., 2009; Sung et al., 2008). In a study by Spörer et al. (2009), third- through sixth-grade students in two German elementary schools were taught to summarize, question, clarify, and make predictions about reading either with peers or an instructor over the course of 9 weeks. Students who received strategy instruction, in either the peer or instructor condition, were better able to ask inferential questions when comprehending text compared with controls, and such skills were evident at a 12-week follow-up. In addition, a subset of 130 sixth-grade Taiwanese students with low and high reading ability were given an 11-week computer-based reading strategy course that taught them to select important content, organize information, integrate new knowledge with former knowledge, self-question, and monitor comprehension. Students at different levels of reading ability in the treatment condition scored higher on an inference measure at posttest. Thus, recent undergraduate-level research on the effect of strategy instruction on inferences is consistent with prior research at the K-12 level and provides further evidence for transfer and longer term retention.

Reading Vocabulary

Vocabulary knowledge is also key for forming a coherent mental model of the text. Understanding the meanings of words makes it possible to form connections between information in different parts of text and to relate what is read to information stored in memory. Providing vocabulary instruction increases comprehension (e.g., Lubliner & Smetana, 2005). Results of previous studies using SEM of longitudinal data on reading development show medium-sized effects of fifth-grade vocabulary on growth in comprehension from fifth to sixth grades (Verhoeven & van Leeuwe, 2008; for correlational evidence, see Blom & Severiens, 2008; Katzir, Lesaux, & KimLaw, 2009; Wood, 2009).

Other Paths

For four other paths in the DIME model, we were not able to locate recent studies meeting the same criteria: the paths from background knowledge to reading strategies and inference, the path from vocabulary to inference, and the path from word reading to comprehension. However, we did locate additional recent correlational studies. These studies strengthen support for the relationship between prior topic knowledge and reading comprehension (e.g., Anmarkrud & Bråten, 2009; Kendeou & van den Broek, 2007; Miller, Cohen, & Wingfield, 2006; Ozuru, Best, Bell, Witherspoon, & McNamara, 2007). Prior topic knowledge is also related to use of the reading comprehension strategies of questioning, note-taking, and use of text structure (Bonner & Holliday, 2006; Surber & Schroeder, 2007; Taboada & Guthrie, 2006; van den Broek & Kendeou, 2008). In addition, prior topic knowledge is related to accurate inferences (Gilabert et al., 2005; van den Broek & Kendeou, 2008). Regression studies demonstrate the small but significant contribution of word reading fluency to reading comprehension even after accounting for other variables, and they demonstrate this contribution into the high school years (e.g., Martínez, Vidal-Abarca, Gil, & Gilabert, 2009; Naceur & Schiefele, 2005; Perfetti et al., 2008; Samuelstuen & Bråten, 2005; Tichá, Espin, & Wayman, 2009; Wood, 2009).

Characteristics of Science Majors and Science Text Related to Text Comprehension

Science Majors

Using data from high school students reading domain-general text, we found a good—but not ideal—fit of the DIME model (Cromley & Azevedo, 2007). Using data from undergraduate

students, we found a good fit of the model (Cromley & Snyder, 2007). By domain-general, we mean that participants were from a diverse set of majors and read texts across a range of academic domains. The literature on comprehension of science text overwhelmingly uses students with nonscience majors as participants (Otero et al., 2002), but these students may differ in important ways from students with science majors—in prior topic knowledge about science, in reading vocabulary (e.g., higher verbal SAT scores), and perhaps in other variables not included in the model (such as various aspects of motivation). Students with science majors and students with nonscience majors may differ not only in mean scores on various predictors, but the interrelationships among the variables in the model may differ for students with science majors versus students with nonscience majors. For example, reading passages chosen for a general audience taking a standardized reading comprehension test might deliberately draw on everyday knowledge and avoid reference to specialized knowledge. On the other hand, passages in a college-level science textbook might deliberately draw on specialized knowledge. The effect of text-specific knowledge might therefore be much higher for the latter text than for the former. In summary, both mean scores on predictors and the size of different effects can affect model fit.

Science Texts

In addition to using domain-general participants, tests of the DIME model have used domain-general texts. Would the DIME model hold equally well with biology students reading biology text? Biology text is characterized by (a) high density of vocabulary that is unfamiliar or that has specific technical meanings that are different from conversational meanings, (b) frequent use of causal and sequential text schemas, and (c) the need to integrate and draw numerous elaborative inferences using knowledge—often only recently acquired—with current segments of text (Diakidoy, Kendeou, & Ioannides, 2003; Otero et al., 2002). Given these unique features of scientific text, if the DIME model holds equally well with scientific text as it does with domain-general text, this would provide evidence for the generalizability of the model beyond domain-general text and could also lead to approaches for improving comprehension among biology students. Furthermore, reading vocabulary, inference, and prior topic knowledge—the variables that differentiate science text from other genres—may make larger contributions to comprehension when testing the DIME model with science-major students reading science text than when testing the model with students reading domain-general text.

Model Fit

In addition to the issues of domain generality and the samples used to test the model, the DIME model showed some weaknesses in model fit. Specifically, there was some suggestion that additional paths should be added to the model. One path identified as having no experimental support was the path from reading vocabulary to reading comprehension strategy use. Although this relationship makes sense—how could a student write a coherent summary of a passage if he or she did not know the meaning of the words?—we were not able to locate any studies that tested this relationship experimentally. Researchers have found statistically significant correlations between knowledge of reading vocabulary and reading comprehension strategy use (Proctor et al., 2007; van Gelderen, Schoonen, Stoel, de Glopper, & Hulstijn, 2007; Willson & Rupley, 1997).

In the present study, we created biology-specific measures of reading comprehension and the other variables in the DIME model, collected data from undergraduate students in an introductory course for biology majors, and tested the fit of the DIME model to these data. Our research questions were as follows:

1. What is the fit of the original DIME model to data from undergraduate-science-major students reading science text?

2. Can we modify the model to obtain a better fit by adding a path from reading vocabulary to reading strategies?

3. What are the predictor variables that make the largest total contribution to reading comprehension from the better fitting model for these students?

Method

Participants

Participants were 737 current or prospective students with biology majors enrolled in a first-semester introductory biology course at a large, moderately selective urban university in the mid-Atlantic region. These represented the 83% of those enrolled in the class who elected to participate in the study in exchange for extra course credit. Students were 18 years of age or older, had already passed first-semester chemistry and math courses, and were concurrently enrolled in second-semester chemistry and math courses. Hence, this was a relatively high-achieving sample of undergraduates. Participants' mean age was 19.77 years ($SD = 2.40$); there were 466 women (63%), 266 men (36%), and five students who did not identify gender. Most of the students were freshmen (50%), with some sophomores (31%), juniors (12%), seniors (4%), and postbaccalaureate students (3%). Participants were racially diverse (37% Asian, 37% White, 14% Black, and 11% mixed race or other races). A substantial minority of participants were first-generation college students; 43% had neither parent with a bachelor's degree or higher. Of those who reported language spoken as a child at home, 47% spoke only English, 20% spoke another language(s) together with English, and 33% spoke a language(s) other than English.

Measures

We followed the same method as in our original test of the DIME model (Cromley & Azevedo, 2007) to develop measures of comprehension and predictors of comprehension. We began with passages about the vertebrate immune system from a popular biology textbook used at both undergraduate and advanced placement biology levels and used as the textbook in the course from which participants were recruited (Campbell & Reece, 2001). These passages about the vertebrate immune system were used to

construct the comprehension measure; all of the other measures also referred to this passage and topic. For example, the prior topic knowledge measure tapped the prerequisite knowledge required to understand the immune system text. Construction of these measures was also informed by results from 40-min think-aloud protocols collected from 91 students in a different section of the same course using the same textbook in Spring 2006 (Cromley, Snyder-Hogan, & Luciw-Dubas, 2010). For example, we strategically embedded some inaccurate inferences that we had observed in the think-aloud protocols as distractors in the inference measure.

Measures included a student demographic questionnaire and topic-specific measures of reading comprehension, prior topic knowledge, inference, reading strategy use, reading vocabulary, and word reading fluency (examples from each measure are given in the Appendix). All measures were administered in a paper-and-pencil format and required silent reading. Other than the demographic and word reading fluency measures, all instruments used multiple-choice items with a stem and four options; correct answers were scored 1, and incorrect answers were scored 0. Measures were pilot tested in Fall 2006 on a sample of students from a different section of the same course, and shortened versions were administered in the current study. Reliability and validity data from the current study are presented with the results.

Reading comprehension. The reading comprehension measure was a 26-item measure that used 10 short passages from the textbook chapter on the immune system and included two line diagrams. The passages explained the purposes, components, and functions of the vertebrate immune system—a topic identified by the course instructor as difficult but manageable for first-semester students. Each passage consisted of one to two paragraphs, was 81–230 words long, was followed by two to four multiple-choice questions, and included a range of question types from lower level recall to higher level application questions. The passages were presented in the same order as in the original textbook chapter.

Prior topic knowledge. The prior topic knowledge measure was a 17-item multiple-choice measure that tested knowledge about and misconceptions about the immune system. The measure was written using nontechnical language to ensure that the knowledge and reading vocabulary measures were tapping different constructs. For example, one item asked the following:

> When a person is exposed to a germ, the immune system: a. Causes the illness; b. Is damaged by the illness; c. Must have been damaged, or else the person would not have gotten the illness; d. Destroys the germs.

See the Results section for analyses to support discriminant validity of the prior topic knowledge and reading vocabulary measures.

Inference. The inference measure was a 10-item multiple-choice measure developed by the researchers based on Hannon and Daneman (2001), which asked students to draw a conclusion from two sentences. The sentences were always true and were either direct quotes from or slightly modified from Campbell and Reece's (2001) textbook chapter. For example, one item presented the following two sentences:

> Animals have three defenses against pathogens and abnormal body cells (which may develop into cancer). AND Two of these lines of defense are not specific to particular infectious agents. THEREFORE: a. Pathogens cause cancer; b. Animals have redundant defenses; c. The non-specific lines of defense are weak; d. Cancer cells require specific defenses.

Option a is a paraphrase of the first sentence; Options b and c represent jumping to conclusions (both statements are true, but the conclusion cannot be drawn from the two sentences provided); and Option d is a correct inference—nonspecific defenses protect against infectious agents only, therefore specific defenses are required for precancerous cells.

Reading comprehension strategy use. The reading comprehension strategy use measure was a 13-item multiple-choice measure based on Kozminsky and Kozminsky (2001). The measure asked students to actually apply specific reading comprehension strategies to passages rather than asking them to self-report on frequency of use of strategies. We chose eight passages from two textbooks: (a) the textbook chapter and (b) passages covering similar material to the comprehension measure from a different undergraduate biology textbook (Starr & McMillan, 2001). We constructed questions that asked students to apply the reading comprehension strategies of summarizing, predicting, self-questioning, prior topic knowledge activation, note-taking (including one diagram drawing question), and coordinating text and diagrams. For example, one prior knowledge activation question asked the following:

> Which of the following would be most useful to know in order to understand the passage? a. Pathogens have evolved defenses against the immune system; b. The body has non-specific defenses as well as specific ones; c. Some kinds of lymphocytes attack all pathogens indiscriminately; d. The surface molecules of each pathogen are unique.

In this example, all four statements were correct, but only one was relevant to the passage. Likewise, one prediction item asked "Which of the following is most likely to follow this passage?" and was followed by four possible predictions. See the Appendix for a sample passage with questions. Each question was scored 1 for selecting the correct answer from the multiple-choice options and 0 for selecting an incorrect answer.

Reading vocabulary. The vocabulary measure was a 19-item multiple-choice measure that tested knowledge of the meanings of both scientific (e.g., antibodies) and nonscientific (e.g., proliferates) terms used in the comprehension passage. The measure used a conventional vocabulary format, presenting a short phrase with one word underlined; participants were asked to choose the option that most closely matched the meaning of the underlined word. As with the other measures, distractors were based on known misunderstandings that we had observed in our prior think-aloud research. We chose vocabulary words that were important for understanding the key concepts in the passage but were not defined in the passage itself. For example, the idea that an antigen elicits a response from immune system cells is critical for understanding the entire chapter, and the word "elicits" is not defined in the chapter. See the Results section for analyses to support discriminant validity of the prior topic knowledge and reading vocabulary measures.

Word reading fluency. The word reading fluency measure was a 1-min timed group-administered maze task based on Naceur and Schiefele (2005). One reason for selecting this type of measure was to allow us to test students in a group rather than having to test

them individually. Tichá et al. (2009) found strong alternate-forms reliability (rs = .84–.88) and validity correlations for eighth-grade students between maze-type curriculum-based word reading fluency measures and scores on the individual, orally administered Woodcock–Johnson III Broad Reading Cluster measure (Woodcock, McGrew, & Mather, 2001) of word reading (rs = .86–.88). In our study, participants silently read a one-page (234-word) passage about the immune system from Starr and McMillan (2001). They were asked to read as quickly as they could while still understanding the material. We ensured that they were reading with comprehension by a embedding a choice of three words every one to two sentences. They were asked to circle the one best word and to continue reading until a timer beeped at exactly 1 min, at which point they made a mark on the page at the last word read. The score was the number of words read in exactly 1 min.

Procedure

Students participated in groups of 5–10 during one 90-min session in our laboratory in Spring 2007 or Spring 2008 using identical procedures. The sessions were offered late in the semester but before the topic of the immune system was covered in the class. After we obtained student informed consent, participants completed a demographics form, and then they completed the word reading fluency (1 min), prior topic knowledge (6 min), reading vocabulary (8 min), reading comprehension (26 min), strategy use (13 min), and inference (13 min) measures. For all of the multiple-choice measures, participants were asked to choose the one best answer for each question and (where applicable) to read the passages before answering questions.

Data analysis. Our primary data analyses used SEM to test the fit of a series of models to the paper-and-pencil data. Our obtained sample size of 737 was larger than the recommended sample size of 460 (5 participants per parameter × 92 parameters) because all students enrolled in the course were offered the opportunity to participate in exchange for extra credit. Because participants completed measures in our laboratory in a single session, we were able to ensure completion of all measures; there were therefore no missing data. Before conducting any analyses, we examined the data for violations of normality, outliers, and nonlinearity; no violations of assumptions were found.

Item parceling. To increase accuracy of parameter estimates, we parceled the items on each measure (except for word reading fluency, which was a single score; Bandalos, 2002). Hall, Snell, and Foust (1999) have provided evidence from a simulation study that parceling can lead to better parameter estimates if there are omitted variables that lead to shared variance among items in a parcel. Parceling can also increase statistical power compared with either path analysis or loading every item from a scale on one factor because fewer parameters are tested (Tempelaar, Gijselaers, van der Loeff, & Nijhuis, 2007). An additional advantage of parceling, as compared with a path analysis using a single summed scale for each construct, includes the ability to estimate errors rather than relying on Cronbach's alpha estimates. In summary, we used a parceling strategy to (a) leave enough degrees of freedom to allow us to test alternative models, (b) increase the accuracy of parameter estimates, and (c) take advantage of partialing out error variance afforded by a full SEM model over path analysis. We split the questions from each measure into random thirds to create three item parcels for each measure. These three parcels were then used as measured indicators of each latent factor.

SEM approach. SEM analyses were carried out with the EQS 6.0 software program (Bentler, 1995) using maximum likelihood estimation. We used the two-step measurement model and the full structural model approach frequently recommended in the literature (Kline, 2005). As is conventional in reporting results of SEM, we report the mediated (indirect) effects as well as the direct effects for each predictor. In SEM terminology, these are meant to represent the contribution of each predictor factor in explaining variance in the criterion factor. We use the term *effect* in its conventional sense, not to imply that we collected the sort of longitudinal or experimental data required to establish causality.

Fit indices. The fit of each model was assessed using the recommendations from Hu and Bentler (1999) for samples of $N > 200$. The fit of nested models was tested using the chi-square test of difference (Kline, 2005). We briefly review these fit indices, for readers who may be less familiar with SEM techniques. All fit indices are based on a comparison of the variances and covariances observed in the data and the variances and covariances that would be predicted if the model was true. When maximum likelihood is used to estimate models, a likelihood function is calculated from the observed and expected variance/covariance matrices. This likelihood function (multiplied by $N - 1$) follows a chi-square distribution, with degrees of freedom equal to the number of free parameters. The chi-square test for a good-fitting model should be nonsignificant (because a good-fitting model leads to tiny residuals). The chi-square statistic is used to calculate the AIC (Akaike information criterion), CFI (confirmatory fit index), and RMSEA (root mean square error of approximation) statistics. There is no numerical benchmark for the AIC; it is a smaller-is-better statistic that can take on negative values. CFIs from a good-fitting model should be greater than .96. The RMSEA has the advantage that it is accompanied by a 90% confidence interval; RMSEAs should be less than .05, or the confidence interval should straddle .05. The fourth fit statistic we report—the standardized root mean residual (SRMR)—is based on the residuals from subtracting the standardized expected variance/covariance matrix from the standardized observed variance/covariance matrix. SRMRs from a good-fitting model should be less than .10 (Hu & Bentler, 1999).

All statistical tests were assessed using an alpha level of $p < .05$. We report obtained p levels for all analyses, except that very small p levels are reported as $p < .001$.

Results

Preliminary Analyses

Descriptive statistics, intercorrelations among all measures, and reliability for all measures are shown in Table 1. All of the measures showed acceptable-to-good reliability.

One concern that has been raised about the DIME model (Cromley & Azevedo, 2007) is that prior topic knowledge and reading vocabulary might not be distinct concepts. Reading vocabulary represents an understanding of the meanings of words, and vocabulary knowledge—including vocabulary for complex concepts—encapsulates knowledge about the world. To ensure that our reading vocabulary and prior topic knowledge measures were tapping separate constructs, we began with a confirmatory factor analysis

Table 1
Descriptive Statistics, Intercorrelations Among All Measures, and Reliability for All Measures

Variable	1	2	3	4	5	6
1. Prior topic knowledge	*.74*					
2. Inference	.45	*.71*				
3. Strategy use	.46	.48	*.67*			
4. Vocabulary	.42	.39	.37	*.76*		
5. Word reading fluency	.25	.16	.24	.25	—	
6. Comprehension	.58	.54	.52	.51	.23	*.80*
M	9.77	8.24	6.39	14.98	144.92	16.39
SD	2.97	2.73	2.22	2.52	38.34	3.71

Note. $N = 737$. Correlations greater than .06 are significant at $p < .05$; correlations greater than .09 are significant at $p < .01$. Reliabilities are shown in italics on the diagonal.

on the vocabulary and background knowledge measures. In this analysis, we fit a model with the vocabulary factor loading only on reading vocabulary subscale scores and with the prior topic knowledge factor loading only on prior topic knowledge subscale scores (shown as Model 7 in Table 2). We then fit a model with both the reading vocabulary and prior topic knowledge factors loading on all subscale scores (all cross-loadings: Model 1). We then removed each cross-loading and tested whether the fit of the model improved (Models 2–6). The most parsimonious model was Model 7—the confirmatory factor analysis model in which each factor loaded only on the subscale scores for the construct we designed it to tap. That model showed an excellent fit, $\chi^2(8) = 6.279$, $p = .616$, $CFI = 1.000$, $SRMR = .016$, $RMSEA < .001$, 90% CI [<.001, .037]. We conclude that the reading vocabulary and prior topic knowledge factors—although correlated—represent distinct and separable factors.

> *Research Question 1:* What is the fit of the original DIME model to data from undergraduate-science-major students reading science text?

We began by fitting a measurement-only model, which is equivalent to fitting a set of confirmatory factor analyses on each factor while simultaneously allowing all factors to correlate with each other. The measurement model showed an excellent fit to the data, $\chi^2(89) = 94.693$, $p = .320$, $AIC = -83.307$, $CFI = .998$, $SRMR = .023$, $RMSEA = .009$, 90% CI [<.001, .023] (see Table 3). This suggests that each factor has a strong relationship to the observed scores that the factor is hypothesized to drive, and therefore fitting a structural model is warranted.

We next tested the fit of the structural model—the original DIME model—which also showed an excellent fit, $\chi^2(92) = 105.065$, $p = .166$, $AIC = -78.935$, $CFI = .995$, $SRMR = .024$, $RMSEA = .012$, 90% CI [<.001, .025], $R^2 = .928$ (see Table 3). This model, however, fit significantly worse than the measurement model, $\Delta\chi^2(3) = 10.372$, $p = .016$, suggesting that the original DIME model may not be an ideal model and could be improved upon.[1] We therefore tested the fit of a modified DIME model, in which we added a path from reading vocabulary to reading comprehension strategies, on the basis of the literature reviewed above.

> *Research Question 2:* Can we modify the model to obtain a better fit by adding a path from vocabulary to reading comprehension strategies?

Using the same model fitting procedures, we tested a modified model including the path from reading vocabulary to strategy use. As noted above, this is an exploratory analysis but one with great theoretical appeal. The modified model also showed an excellent fit to the data, $\chi^2(91) = 99.974$, $p = .244$, $CFI = .996$, $SRMR = .023$, $RMSEA = .012$, 90% CI [<.001, .024], $R^2 = .928$. This model showed a significantly better fit than the measurement model, $\Delta\chi^2(2) = 5.281$, $p = .070$, and showed a significantly better fit than the original DIME model, $\Delta\chi^2(1) = 5.091$, $p = .020$. This suggests that adding the path from vocabulary to reading comprehension strategies results in a better model. The finding of a significantly better fit for the modified DIME model suggests that it is a plausible model for explaining these data collected from biology-major students reading biology-specific text. It is important to keep in mind that this is not the only plausible or possible model for explaining these data, simply that it fit the data well.

In the final model, six of the 10 paths in the model were statistically significantly different from zero, and all three correlations were significant (see Figure 2). The direct effects of reading strategies and word reading on comprehension were nonsignificant, as were the direct effects of prior topic knowledge and vocabulary on inference.

> *Research Question 3:* What are the predictor variables that make the largest total contribution to reading comprehension from the better fitting model for these students?

For the best fitting modified DIME model, the total contribution of each predictor to comprehension is shown in Table 4. These total effects are calculated by adding the direct effect of each predictor—equivalent to a beta weight in a regression—on comprehension to the indirect effects of each predictor on comprehension (see Loehlin, 2004, for path tracing rules used to calculate these indirect effects). Prior topic knowledge had by far the largest effect on comprehension, due both to its significant direct effect (standardized path loading = .345) and also three indirect effects

[1] We attempted to fit a model that regressed reading comprehension on word reading and a listening-comprehension-like factor that loaded on the other predictor variables. Because of the small contribution of our word reading fluency measure to comprehension, we were not able to get the estimates to converge because the listening-comprehension-like factor was essentially the same as reading comprehension.

Table 2
Fit Indices for a Series of Models to Show Separability of Prior Topic Knowledge and Vocabulary Factors

Model	χ^2 (*df*)	Change in χ^2 (*df*)	Significance of change
Model 1: Factors cross-load on all scores	1.046 (2)		
Model 2: Factors load on all scores but one (Factor 1 → Variable 6)	1.046 (3)	<0.001 (1)	1.000
Model 3: Model 2 with one path dropped (Factor 1 → Variable 5)	3.624 (4)	2.578 (1)	.108
Model 4: Model 3 with one path dropped (Factor 1 → Variable 4)	3.678 (5)	0.054 (1)	.816
Model 5: Model 4 with one path dropped (Factor 2 → Variable 3)	3.678 (6)	<0.001 (1)	1.000
Model 6: Model 5 with one path dropped (Factor 2 → Variable 2)	6.276 (7)	2.598 (1)	.107
Model 7: Factor loads only on scores intended to underlie that factor	6.279 (8)	0.003 (1)	.956

via reading strategies (statistically significant) and/or inference (nonsignificant), yielding total indirect effects = .281. The next largest contributions to comprehension were made by reading vocabulary, inference, and reading strategies. Reading vocabulary had both a significant direct effect (path loading = .278) and a significant indirect effect (path loading = .119) via inference and strategies. Inference had a significant direct effect (path loading = .394) on comprehension, and reading strategies had a nonsignificant direct effect (path loading = .081) and a significant indirect effect via inference (path loading = .323). Word reading fluency had a negligibly small effect that was nonsignificantly different from zero (path loading = −.041).

Discussion

In the series of models we tested, the original DIME model showed a significantly worse fit to the data than the measurement model. This led us to test a more exploratory, but theoretically justified, modification of the model. The modified DIME model, which included a path from reading vocabulary to strategy use, showed a statistically significantly better fit than either the measurement model or the original model. In the final model, six of the 10 paths and all of the correlations were significantly different from zero, with the largest total effect from prior topic knowledge, followed by inference and reading vocabulary.

In the final model, there were medium-sized indirect effects of prior topic knowledge (via reading comprehension strategies and inference) on reading comprehension. There were medium-sized indirect effects of reading comprehension strategies (via inference) on reading comprehension. Finally, there were medium-sized indirect effects of vocabulary (via reading comprehension strategies and inference) on reading comprehension. Effects of reading comprehension strategies were almost entirely indirect, via inference. Word reading showed a nonsignificant direct effect.

These results are consistent with our prior tests of the DIME model (Cromley & Azevedo, 2007; Cromley & Snyder, 2007), in the sense that the original DIME model showed an excellent fit to the data. The current results diverge from our prior findings with domain-general text, in that they suggest an unusually strong role for topic knowledge in the comprehension of biology text, where previously we had found the strongest role for vocabulary in the comprehension of domain-general text. Our results suggest that biology knowledge has a strong statistically significant relationship with both strategy use and inference in biology text, as well as a direct statistically significant relationship with comprehension, even after controlling for all other predictors. The importance of topic knowledge for the comprehension of scientific text is well supported by a number of prior studies (Bonner & Holliday, 2006; Gilabert et al., 2005; Kendeou & van den Broek, 2007; Miller et al., 2006; Ozuru et al., 2007; Surber & Schroeder, 2007; Taboada & Guthrie, 2006; van den Broek & Kendeou, 2008), but our study shows its importance even after controlling for strategy use, inference, vocabulary, and word reading.

Also consistent with our prior test of the DIME model, inference had a significant direct effect on comprehension; this effect was larger than we had found previously (Cromley & Azevedo, 2007; .39 vs. .19). One possible reason for the stronger effects of both prior topic knowledge and inference on comprehension in the current study is the nature of biology text. First, authors of undergraduate textbooks may assume greater and more specialized knowledge of biological concepts from students in an introductory college course than do authors of other types of text. Text selections used in general reading comprehension tests, on the other

Table 3
Indicators of Fit for the Original and Modified Direct and Inferential Mediation (DIME) Models

Fit index	Measurement model	Original DIME model	Modified DIME model
χ^2 (*df*)	94.693_a (89)	105.065_b (92)	99.974_a (91)
AIC	−83.307	−78.935	−82.026
CFI	.998	.995	.996
SRMR	.023	.024	.023
RMSEA [90% CI]	.009 [<.001, .023]	.014 [<.001, .025]	.012 [<.001, .024]
R^2		.929	.928

Note. Different subscripts indicate statistically significant differences ($p < .05$). AIC = Akaike information criterion; CFI = confirmatory fit index; SRMR = standardized root mean residual; RMSEA = root mean square error of approximation; CI = confidence interval.

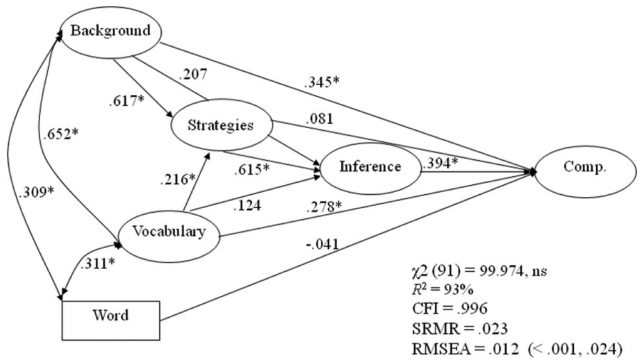

Figure 2. Path coefficients and fit indices for the final modified direct and inferential mediation model. Measured indicators and error terms were omitted for clarity. Comp. = comprehension; CFI = confirmatory fit index; SRMR = standardized root mean residual; RMSEA = root mean square error of approximation. $^* p < .05$.

hand, may be deliberately chosen to minimize the need for any specialized topic knowledge. Second, undergraduate-level textbooks, like much of scientific text, develop concepts in a less coherent way than other types of text (i.e., using fewer explicit connectives from between sentences and concepts; Goldman & Bisanz, 2002); they therefore require students to draw more inferences.

Contrary to our expectations, vocabulary had a somewhat smaller direct effect on comprehension than we had found previously with high school students (Cromley & Azevedo, 2007; .27 vs. .37). This could be due to any number of differences between the studies—previously we had used a standardized measure of general reading vocabulary, whereas in the present study we used a text-specific measure. Presumably, general measures of vocabulary are picking up on variance from other, unmeasured variables. As noted previously, there are age, domain, and measurement differences between the two studies. In our earlier test of the DIME model, we used a path analysis approach, whereas in the present study, we used SEM. In addition, we included an additional indirect effect of reading vocabulary on reading comprehension strategy use. Students who use mental resources to figure out the meaning of new vocabulary might be at a disadvantage compared with those who already know the meaning of the word and can immediately proceed to using a strategy such as summarizing or generating questions.

Some insight into the relationship between vocabulary and comprehension in this type of text might be provided by the think-aloud study we conducted with a different sample from this same class reading from the same textbook. In our prior think-aloud research, we were surprised to find no relationship between verbalizations of vocabulary difficulty and gains in knowledge from before reading until after reading. It appeared that the textbook did indeed have a heavy vocabulary load (Goldman & Bisanz, 2002) and that many students lacked the relevant vocabulary, but the text provided explicit definitions for many of these terms. Students who learned more from the passage were better able to learn new vocabulary while reading, even though they expressed as many difficulties when they first encountered a new word.

Consistent with our prior test of the DIME model, in the present study, reading comprehension strategies had their effect indirectly via inference, and they showed no statistically significant direct effect on comprehension. This finding is at odds with numerous recommendations in the reading instruction literature that cognitive strategy instruction directly increases reading comprehension (e.g., Kamil et al., 2008).

Another difference from prior results is the nonsignificant effect of word reading fluency on reading comprehension. Other research with high school and undergraduate students (Artelt, Schiefele, & Schneider, 2001; Naceur & Schiefele, 2005; Perfetti et al., 2008; Samuelstuen & Bråten, 2005) continues to show a significant contribution of word reading fluency to comprehension. However, we also note that word reading fluency was highly collinear with the other predictors. This means that studies that do not include this set of predictors are capturing some of the variance of the missing predictors in word reading fluency (and vice versa). One possible explanation for our findings is the population we studied: These were relatively high-achieving students who had already passed one semester of demanding chemistry and math courses and were simultaneously enrolled in biology, chemistry, and math courses at the time of the study. For this relatively high-achieving population of students, there was not enormous variability in word reading fluency; other research with undergraduate populations has drawn on students with documented reading problems who show much more variability on word reading (e.g., Bell & Perfetti, 1994). In addition, our prior think-aloud study with students from this course showed no relationship between number of words read and quality of free recall after reading, and there was not any relationship between word reading errors and quality of free recall. It seems that college students can read this text slowly and mispronounce many technical terms but still learn a great deal from the text.

Limitations

We note a number of limitations in our measures and limitations due to other aspects of the study design. Reliability of the measures was only acceptable-to-good (Cronbach's αs = .67–.80), with no reliability available for the word reading fluency measure. Cronbach's alpha tests the uni-dimensionality of measures, and it is likely that inference and strategy use are multidimensional. For example, students may be strong at summarizing or predicting but may be weak at generating questions or using a concept map.

Table 4
Standardized Direct and Indirect Effects of Predictors on Reading Comprehension in the Modified Direct and Inferential Mediation Model

Variable	Direct	Indirect	Total
Prior topic knowledge	.345*	.281*	.626*
Inference	.394*	—	.394*
Strategies	.081	.242*	.323*
Vocabulary	.278*	.119*	.397*
Word reading	−.041	—	−.041

Note. A dash indicates that the path is not included in the model.
$^* p < .05$ (with a one-tailed z test).

Furthermore, we did not test strategies specific to scientific text structure, such as nominalization (Unsworth, 1999). If we had used a larger number and variety of items in these measures, we might have been able to create more reliable secondary factors, which would have increased measurement precision. With regard to word reading fluency, there are complex tradeoffs in efficiently conducting large-scale studies such as this one. Individually administered measures of word reading fluency may be more accurate and do allow for the calculation of split-half reliabilities compared with the group-administered measures used here, but they require a larger research staff than we had available to accommodate the 5–10 students we tested at one time. We do note that using an SEM analytical approach accounts statistically for the unreliability of the measures. We also note that we were not able to use data collected with multiple methods (e.g., observations of study strategies, teacher reports of comprehension) as indicators of each latent construct. This means that factors capture some of the shared variance among the parcels, which all use multiple-choice, paper-and-pencil, student-provided data.

Other limitations of the study relate to the sample, our SEM approach, and our domain-specific approach. Our sample was drawn from only one course at one institution, although this was a large course with a very diverse enrollment. There are possible risks of item parceling in SEM, especially bias in parameters and poor model fit. However, parameter estimates are not dramatically different from prior research, our confirmatory factor analysis on background knowledge and vocabulary is a check against multidimensionality, and the final model showed an excellent fit to the data. We see certain advantages to having students enrolled in science majors complete domain-specific measures, where for example vocabulary questions test the vocabulary used in the reading passage. This tradeoff, however, reduces the generalizability of our findings—we have no information about how the model would fit data from, for example, history-major students reading history text. These results could be specific to science text, biology text, or perhaps only immune system text.

Conclusion

We have shown that the DIME model, which we originally tested with domain-general text on students who were not specializing in any particular domain, also shows an excellent fit with domain-specific biology text read by students with biology majors. This suggests that the DIME model can be generalized beyond domain-general participants and domain-general texts to biology students reading biology text, and the model can still show an excellent fit. Across domains and majors, there are some similarities and differences in the findings. In all of the studies, four predictors—prior topic knowledge, reading vocabulary, strategies, and inference—had significant effects on comprehension; only for high school students was word reading a significant predictor. Strategies had a significant direct effect only for the undergraduate students reading domain-general texts. Whereas previously we had found the largest effects for vocabulary with high school students and undergraduate students reading domain-general text, it was background knowledge that showed the largest effects for the biology students in the current study. We argue that the strong effects of background knowledge and inference in the DIME model when it is tested with biology students reading biology text speaks to the different demands that this type of text places on readers (Goldman & Bisanz, 2002).

In this study, we have validated the DIME model with domain-specific biology text read by students with biology majors but with one modification that made a large and significant improvement in the fit of the model—adding a direct effect of vocabulary on reading comprehension strategies. This finding suggests that students with more vocabulary knowledge—both general academic vocabulary and scientific vocabulary—also show more accurate use of strategies and higher comprehension. Our results also suggest that students with more scientific background knowledge have better reading comprehension of science text because of both a direct effect of background knowledge and indirect effects via strategies and inference.

We see three specific educational implications of these findings: One is that students could be encouraged to build their background knowledge as much as possible before they need to comprehend science texts such as these (e.g., entering freshmen with lower advanced placement examination scores could be encouraged to take the introductory courses because this is likely to build their knowledge and thereby increase their comprehension in later courses). Second, instructors might consider either building basic knowledge before students read a chapter (e.g., with a mini-lesson at the end of each lecture or an advance organizer to prepare students for the next reading) and/or confronting misconceptions before students read a chapter (e.g., the immune system fights off infections; it is rarely damaged by infections [with HIV/AIDS being an important exception]). Third, students with low background knowledge might not benefit as much from strategy and/or inference interventions (e.g., self-explanation; Roy & Chi, 2005) as students with more background knowledge. Finally, even if students have background knowledge, they may need to be reminded to activate it and use it while reading.

Theoretically, these results provide support for the DIME model with a new domain-specific sample. All three tests of the DIME model suggest that reading comprehension strategies have only a significant indirect effect via inference on reading comprehension. Results with this biology-specific sample and text are mostly consistent with analyses of scientific text (Goldman & Bisanz, 2002)—scientific text places particularly heavy demands on prior topic knowledge and inference. The one departure from prior findings regarding scientific text is that the vocabulary-considerate style of this particular biology textbook may counteract its heavy vocabulary demands. In future research, we would like to test the model with longitudinal data to see how the relative sizes of the effects change with development and to consider the role of motivation in the model.

References

Ainsworth, S., & Burcham, S. (2007). The impact of text coherence on learning by self-explanation. *Learning and Instruction, 17,* 286–303.

Anmarkrud, Ø., & Bråten, I. (2009). Motivation for reading comprehension. *Learning and Individual Differences, 19,* 252–256.

Artelt, C., Schiefele, U., & Schneider, W. (2001). Predictors of reading literacy. *European Journal of Psychology of Education, 16,* 363–383.

Bandalos, D. L. (2002). The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. *Structural Equation Modeling, 9,* 78–102.

Bell, L. C., & Perfetti, C. A. (1994). Reading skill: Some adult comparisons. *Journal of Educational Psychology, 86,* 244–255.

Bentler, P. (1995). *EQS structural equation program manual.* Encino, CA: Multivariate Software.

Best, R., Rowe, M., Ozuru, Y., & McNamara, D. (2005). Deep-level comprehension of science texts: The role of the reader and the text. *Topics in Language Disorders, 25,* 65–83.

Blom, S., & Severiens, S. (2008). Engagement in self-regulated deep learning of successful immigrant and non-immigrant students in inner-city schools. *European Journal of Psychology of Education, 23,* 41–58.

Bonner, J., & Holliday, W. (2006). How college science students engage in note-taking strategies. *Journal of Research in Science Teaching, 43,* 786–818.

Broer, N. A., Aarnoutse, C. A. J., Kieviet, F. K., & van Leeuwe, J. F. J. (2002). The effect of instructing the structural aspect of texts. *Educational Studies, 28,* 213–238.

Callender, A., & McDaniel, M. (2007). The benefits of embedded question adjuncts for low and high structure builders. *Journal of Educational Psychology, 99,* 339–348.

Campbell, N. A., & Reece, J. B. (2001). *Biology* (6th ed.). San Francisco, CA: Benjamin Cummings.

Cromley, J. G. (2009). Reading achievement and science proficiency: International comparisons from the Programme on International Student Assessment. *Reading Psychology, 30,* 89–116. doi:10.1080/02702710802274903

Cromley, J., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology, 99,* 311–325.

Cromley, J. G., & Snyder, L. E. (2007). *Testing the fit of the DIME model of reading comprehension with undergraduate students.* Paper presented at the annual conference of the American Educational Research Association, Chicago, IL.

Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Cognitive processes in complex science text and diagrams. *Contemporary Educational Psychology, 35,* 59–74. doi:10.1016/j.cedpsych.2009.10.002

De Corte, E., Verschaffel, L., & Van De Ven, A. (2001). Improving text comprehension strategies in upper primary school children: A design experiment. *British Journal of Educational Psychology, 71,* 531–559.

Diakidoy, I.-A. N., Kendeou, P., & Ioannides, C. (2003). Reading about energy: The effects of text structure in science learning and conceptual change. *Contemporary Educational Psychology, 28,* 335–356.

Gilabert, R., Martinez, G., & Vidal-Abarca, E. (2005). Some good texts are always better: Text revision to foster inferences of readers with high and low prior background knowledge. *Learning and Instruction, 15,* 45–68.

Goldman, S. R., & Bisanz, G. (2002). Toward a functional analysis of scientific genres: Implications for understanding and learning processes. In J. Otero, J. León, & A. Graesser (Eds.), *The psychology of science text comprehension* (pp. 19–50). Mahwah, NJ: Erlbaum.

Gunn, T. M. (2008). The effects of questioning on text processing. *Reading Psychology, 29,* 405–442.

Hall, R. J., Snell, A. F., & Foust, M. S. (1999). Item parceling strategies in SEM: Investigating the subtle effects of unmodeled secondary constructs. *Organizational Research Methods, 2,* 233–256. doi:10.1177/109442819923002

Hannon, B., & Daneman, M. (2001). A new tool for measuring and understanding individual differences in the component processes of reading comprehension. *Journal of Educational Psychology, 93,* 103–128.

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research, 66,* 99–136.

Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55.

Kamil, M., Borman, G., Dole, J., Kral, C., Salinger, T., & Torgesen, J. (2008). *Improving adolescent literacy: Effective classroom and intervention practices: A practice guide* (NCEE No. 2008-4027). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc

Katzir, T., Lesaux, N. K., & KimLaw, Y.-S. (2009). The role of reading self-concept and home literacy practices in fourth grade reading comprehension. *Reading and Writing, 22,* 261–276.

Kendeou, P., & van den Broek, P. (2007). The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition, 35,* 1567–1577.

Kim, J., & White, T. (2008). Scaffolding voluntary summer reading for children in Grades 3 to 5: An experimental study. *Scientific Studies of Reading, 12,* 1–23.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* New York, NY: Cambridge University Press.

Kline, R. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.

Kozminsky, E., & Kozminsky, L. (2001). How do general knowledge and reading strategies ability relate to reading comprehension of high school students at different educational levels? *Journal of Research in Reading, 24,* 187–204.

Lau, K., & Chan, D. (2007). The effects of cognitive strategy instruction on Chinese reading comprehension among Hong Kong low achieving students. *Reading and Writing: An Interdisciplinary Journal, 20,* 833–857.

Law, Y., Chan, C. K. K., & Sachs, J. (2008). Beliefs about learning, self-regulated strategies and text comprehension among Chinese children. *British Journal of Educational Psychology, 78,* 51–73.

Loehlin, J. C. (2004). *Latent variable models: An introduction to factor, path, and structural equation analysis* (4th ed.). Mahwah, NJ: Erlbaum.

Lubliner, S., & Smetana, L. (2005). The effects of comprehensive vocabulary instruction on Title I students' metacognitive word-learning skills and reading comprehension. *Journal of Literacy Research, 37,* 163–200.

Martínez, T., Vidal-Abarca, E., Gil, L., & Gilabert, R. (2009). On-line assessment of comprehension processes. *The Spanish Journal of Psychology, 12,* 308–319.

McCrudden, M. T., Schraw, G., & Lehman, S. (2009). The use of adjunct displays to facilitate comprehension of causal relationships in expository text. *Instructional Science, 37,* 65–86.

McDaniel, M. A., Howard, D. C., & Einstein, G. O. (2009). The read-recite-review study strategy. *Psychological Science, 20,* 516–522.

McKeown, M. G., Beck, I. L., & Blake, R. G. K. (2009). Rethinking reading comprehension instruction: A comparison of instruction for strategies and content approaches. *Reading Research Quarterly, 44,* 218–253.

McNamara, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes, 38,* 1–30.

Miller, L. S., Cohen, J., & Wingfield, A. (2006). Contextual knowledge reduces demands on working memory during reading. *Memory & Cognition, 34,* 1355–1367.

Naceur, A., & Schiefele, U. (2005). Motivation and learning—The role of interest in construction of representation of text and long-term retention: Inter- and intraindividual analyses. *European Journal of Psychology of Education, 20,* 155–170.

Nation, K., & Angell, P. (2006). Learning to read and learning to comprehend. *London Review of Education, 4,* 77–87.

National Research Council. (2007). *Rising above the gathering storm: Energizing and employing America for a brighter future.* Washington, DC: Author.

O'Reilly, T., & McNamara, D. (2007a). The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional "high-stakes" measures of high school students' science achievement. *American Educational Research Journal, 44,* 161–196.

O'Reilly, T., & McNamara, D. S. (2007b). Reversing the reverse cohesion effect: Good texts can be better for strategic, high-knowledge readers. *Discourse Processes, 43,* 121–152.

Otero, J., León, J., & Graesser, A. (Eds.). (2002). *The psychology of science text comprehension.* Mahwah, NJ: Erlbaum.

Ozgungor, S., & Guthrie, J. (2004). Interactions among elaborative interrogation, knowledge, and interest in the process of constructing knowledge from text. *Journal of Educational Psychology, 96,* 437–443.

Ozuru, Y., Best, R., Bell, C., Witherspoon, A., & McNamara, D. S. (2007). Influence of question format and text availability on the assessment of expository text comprehension. *Cognition and Instruction, 25,* 399–438.

Ozuru, Y., Dempsey, K., & McNamara, D. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction, 19,* 228–242.

Paris, S., & Stahl, S. (2005). *Children's reading comprehension and assessment.* Mahwah, NJ: Erlbaum.

Perfetti, C., Yang, C., & Schmalhofer, F. (2008). Comprehension skill and word-to-text integration processes. *Applied Cognitive Psychology, 22,* 303–318.

Proctor, P., Dalton, B., & Grisham, D. (2007). Scaffolding English language learners and struggling readers in a universal literacy environment with embedded strategy instruction and vocabulary support. *Journal of Literacy Research, 39,* 71–93.

Roy, M., & Chi, M. H. (2005). The self-explanation principle in multimedia learning. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 271–286). New York, NY: Cambridge University Press.

Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. *Educational Psychologist, 24,* 113–142.

Samuelstuen, M., & Bråten, I. (2005). Decoding, knowledge, and strategies in comprehension of expository text. *Scandinavian Journal of Psychology, 46,* 107–117.

Savolainen, H., Ahonen, T., Aro, M., Tolvanen, A., & Holopainen, L. (2008). Reading comprehension, word reading and spelling as predictors of school achievement and choice of secondary education. *Learning and Instruction, 18,* 201–210.

Souvignier, E., & Mokhlesgerami, J. (2006). Using self-regulation as a framework for implementing strategy instruction to foster reading comprehension. *Learning and Instruction, 16,* 57–71.

Spörer, N., Brunstein, J. C., & Kieschke, U. (2009). Improving students' reading comprehension skills: Effects of strategy instruction and reciprocal teaching. *Learning and Instruction, 19,* 272–286.

Starr, C., & McMillan, B. (2001). *Human biology* (4th ed.). Pacific Grove, CA: Wadsworth.

Sung, Y., Chang, K., & Huang, J. (2008). Improving children's reading comprehension and use of strategies through computer-based strategy training. *Computers in Human Behavior, 24,* 1552–1571.

Surber, J. R., & Schroeder, M. (2007). Effect of prior domain knowledge and headings on processing of informative text. *Contemporary Educational Psychology, 32,* 485–498.

Taboada, A., & Guthrie, J. (2006). Contributions of student questioning and prior knowledge to construction of knowledge from reading information text. *Journal of Literacy Research, 38,* 1–35.

Tempelaar, D., Gijselaers, W., van der Loeff, S. S., & Nijhuis, J. (2007). A structural equation model analyzing the relationship of student personality factors and achievement motivations, in a range of academic subjects. *Contemporary Educational Psychology, 32,* 105–131.

Tichá, R., Espin, C. A., & Wayman, M. M. (2009). Reading progress monitoring for secondary-school students: Reliability, validity, and sensitivity to growth of reading-aloud and maze-selection measures. *Learning Disabilities Research & Practice, 24,* 132–142.

Unsworth, L. (1999). Developing critical understanding of the specialised language of school science and history texts: A functional grammatical perspective. *Journal of Adolescent and Adult Literacy, 42,* 508–521.

van den Broek, P., & Kendeou, P. (2008). Cognitive processes in comprehension of science texts: The role of co-activation in confronting misconceptions. *Applied Cognitive Psychology, 22,* 335–351.

van Gelderen, A., Schoonen, R., Stoel, R., de Glopper, K., & Hulstijn, J. (2007). Development of adolescent reading comprehension in Language 1 and Language 2: A longitudinal analysis of constituent components. *Journal of Educational Psychology, 99,* 477–491.

Verhoeven, L., & van Leeuwe, J. (2008). Prediction of the development of reading comprehension: A longitudinal study. *Applied Cognitive Psychology, 22,* 407–423.

Wigfield, A., Guthrie, J. T., Perencevich, K. C., Taboada, A., Klauda, S. L., McRae, A., & Barbosa, P. (2008). Role of reading engagement in mediating effects of reading comprehension instruction on reading outcomes. *Psychology in the Schools, 45,* 432–445.

Willson, V., & Rupley, W. (1997). A structural equation model for reading comprehension based on background, phonemic, and strategy knowledge. *Scientific Studies of Reading, 1,* 45–63.

Wood, D. E. (2009). Modeling the relationships between cognitive and reading measures in third and fourth grade children. *Journal of Psychoeducational Assessment, 27,* 96–112.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock–Johnson III.* Rolling Meadows, IL: Riverside Publishing.

Appendix

Sample Items From the Prior Topic Knowledge, Inference, Reading Strategy, Reading Vocabulary, and Reading Comprehension Measures

Prior Topic Knowledge

4. A fever is
   - a. An immune response
   - b. A sign that the immune system has been damaged
   - c. Heat from bacteria multiplying
   - d. A disease

Inference

2. The body's internal non-specific defense system depends on phagocytosis, where certain types of white blood cells ingest pathogens.
AND
Macrophages are very effective, large, long-living phagocytes.
THEREFORE
   - a. Macrophages are white blood cells
   - b. Non-specific defenses are very effective
   - c. All white blood cells ingest pathogens
   - d. The immune system operates by phagocytizing pathogens

Reading Strategy (from Starr & McMillan, 2001, p. 162)

One key feature of the immune system is immunological specificity. This is the ability of certain kinds of lymphocytes to zero in on specific pathogens and eliminate them. Any molecular feature that triggers the formation of lymphocyte armies and is their target is an antigen. The most important antigens are certain proteins at the surface of pathogens or tumor cells, or ones that are unbound but toxic. As you will see, lymphocytes can recognize nonself because they have receptors that bind to such targeted "foreign" features.

5. Which of the following would be most useful to know in order to understand the passage?
   - a. Pathogens have evolved defenses against the immune system
   - b. The body has non-specific defenses as well as specific ones
   - c. Some kinds of lymphocytes attack all pathogens indiscriminately
   - d. The surface molecules of each pathogen are unique

Reading Vocabulary

5. to proliferate
   - a. to make a profile
   - b. to encourage life or growth
   - c. to make low-class
   - d. to make or build

Word Reading Fluency (Sample Sentence; from Starr & McMillan, 2001, p. 168)

The second feature is immunological memory—that is, some
   immunity
   lymphocytes that form during an initial confrontation are set aside for a future
   features
battle with the same pathogen.

(*Appendix continues*)

Reading Comprehension (from Campbell & Reece, 2001, p. 907)

Lymphocyte Development Gives Rise to an Immune System That Distinguishes Self From Nonself

B cells and T cells mature in the bone marrow and thymus. Lymphocytes bearing receptors specific for molecules already present in the body are either rendered nonfunctional or destroyed by apoptosis (programmed cell death). Thus the body normally has no mature lymphocytes that react against self components: The immune system exhibits the critical feature of self-tolerance. Failure of self-tolerance can lead to autoimmune diseases such as multiple sclerosis.

16. Lymphocytes that are selected by your own cells are destroyed by your body. Why is this the case?

a. Because the body has no mature lymphocytes
b. Because these lymphocytes would start an immune response and destroy your own cells
c. This only happens when a person has an autoimmune disorder
d. This is a way for the body to get rid of old lymphocytes

Received February 9, 2009
Revision received January 11, 2010
Accepted March 10, 2010 ■